CLAIMS

What is claimed is:

1.   A method for determining similarity between a plurality of musical works comprising the steps of:

    obtaining respective digitized audio files of the plurality of musical works;

    for each musical work in the plurality, forming (i) a spectral representation from the corresponding audio file and (ii) a rhythmic beat representation from the corresponding audio file;

    for a given musical work of interest:

        (a)   comparing its spectral representation to the spectral representations of the musical works in the plurality;

        (b)   comparing its rhythmic beat representation to the rhythmic beat representations of the musical works in the plurality; and

        (c)   summing, including respective weighting of results of the comparisons in (a) and (b), said summed results providing an indication of which musical works in the plurality are similar to the given musical work of interest.

2.   The method of Claim 1, wherein forming a spectral representation includes dividing the corresponding audio file into a plurality of frames.

3.   The method of Claim 2, further comprising converting each frame to a spectral representation to obtain a plurality of spectral representations for the audio file.

4.   The method of Claim 3, wherein the spectral representation includes a vector of Mel-frequency cepstral coefficients.

5. The method of Claim 3, wherein each spectral representation includes a plurality of Mel-frequency cepstral coefficients.

6. The method of Claim 2, further comprising performing a windowing function on each frame.

5 7. The method of Claim 2, further comprising applying a Hamming window on each frame.

8. The method of Claim 2, further comprising applying a pre-emphasis on each frame.

9. The method of Claim 8, further comprising subjecting data from each frame to a Fast Fourier Transform function to obtain a frequency domain signal for each frame.

10 10. The method of Claim 9, further comprising warping a log amplitude of each frequency signal to a Mel-frequency scale.

11. The method of Claim 10, further comprising subjecting the warped frequency function to a second Fast Fourier Transform to obtain a parameter set of Mel-frequency cepstral coefficients.

15 12. The method of Claim 10, further comprising subjecting the frequency domain signal for each frame to a set of triangular filters to obtain a plurality of Mel-frequency spaced components.

13. The method of Claim 12, further comprising subjecting the Mel-frequency spaced components to a discrete cosine transform function to obtain a plurality of Mel-
20 frequency cepstral coefficients.

14.   The method of Claim 3, further comprising clustering the spectral representations of the audio file to obtain a spectral signature for the audio file.

15.   The method of Claim 14, further comprising comparing the spectral signatures of two different audio files using an Earth Mover's Distance.

5   16.   The method of Claim 1, wherein forming a rhythmic beat representation includes dividing the corresponding audio file into a plurality of frames.

17.   The method of Claim 16, further comprising converting each frame to a spectral representation to obtain a plurality of spectral representations for the audio file.

18.   The method of Claim 17, wherein each spectral representation includes a plurality of
10   Mel-frequency cepstral coefficients.

19.   The method of Claim 17, further comprising computing a similarity matrix for the audio file.

20.   The method of Claim 19, further comprising computing a beat spectrogram for the audio file.

15   21.   The method of Claim 20, further comprising constructing a histogram of the beat spectrogram.

22.   The method of Claim 21, further comprising normalizing the histogram to account for the total number of frames of the audio file.

23.   The method of Claim 22, further comprising calculating a distance between a pair of
20   histograms.

24.   The method of Claim 23, wherein calculating the distance includes calculating the closest distance between the pair of histograms.

25.   The method of Claim 24, wherein the closest distance is the minimum of the sum of absolute differences between bins of each histogram calculated over a range of scalings of each histogram.

5

26.   The method of Claim 25, further comprising applying a function to each histogram to weight certain bins.

27.   The method of Claim 26, further comprising scaling each histogram at least twice to allow for slight differences between musical works.

10   28.   The method of Claim 27, wherein for each scale factor, one histogram is resampled by a factor and compared to the unscaled histogram.

29.   The method of Claim 1, further comprising generating a set of similar musical works.

30.   The method of Claim 29, further comprising visually displaying the musical works in a manner illustrating relative similarities or dissimilarities.

15   31.   The method of Claim 1, further comprising calculating a relative distance between each pair of musical works.

32.   The method of Claim 31, further comprising constructing a matrix of song similarity based on the relative distance.

33.   The method of Claim 32, further comprising performing a Multi-dimensional scaling
20   on the matrix to obtain coordinates in K-dimensional space for each musical work, one coordinate per song.

34. The method of Claim 33, further comprising plotting the coordinates.

35. A method for determining similarity between a plurality of musical works comprising:

obtaining respective digitized audio files of the plurality of musical works;

for each musical work, forming at least two different representations from the corresponding audio file, the different representations representing respective different aspects of the musical work;

for a given musical work of interest:

(a) comparing one of its two different representations to respective ones of the two different representations of the musical works in the plurality;

(b) comparing the other of the two different representations of the given musical work to respective other ones of the two different representations of the musical works in the plurality; and

(c) summing results of the comparisons in (a) and (b), said summed results providing a quantitative indication of which musical works in the plurality are similar to the given musical work of interest.

36. The method of Claim 35, wherein the step of forming at least two different representations includes forming a spectral representation and a beat representation for each musical work, the spectral representation representing instrumentation of the musical work and the beat representation representing rhythmic frequencies of the musical work.

37. The method of Claim 35, further comprising the step of preprocessing the audio files before forming the different representations for each musical work.

38. The method of Claim 37, wherein the step of preprocessing includes omitting relatively long pauses.

39.   The method of Claim 35, further comprising providing a respective reliability measure associated with each representation.

40.   The method of Claim 35, wherein the step of summing includes weighting results of the comparisons as a function of reliability measures of the representations compared.

5   41.   The method of Claim 35, further comprising visually displaying the plurality of musical works in a manner illustrating relative similarities and dissimilarities among the plurality.

42.   A method of processing a database of musical works, comprising:

   obtaining a digitized audio file for each musical work;

10   obtaining a spectral representation for each audio file;

   obtaining a rhythmic beat representation for each audio file;

   for each musical work, summing the spectral representation and the rhythmic beat representation; and

   determining a similarity of the plurality of musical works based on the summed

15   results.

43.   The method of Claim 42, further comprising weighting the summed results.

44.   The method of Claim 42, further comprising visually displaying the plurality of musical works including indicating the determined similarities.

45.   A computer program product for determining similarity between a plurality of musical

20   works, the computer program product including a computer usable medium having computer readable code thereon, including program code which:

   obtains respective digitized audio files of the plurality, and forms (i) a spectral representation from the corresponding audio file and (ii) a rhythmic beat representation from the corresponding audio file; and

for a given musical work of interest:

(a)   compares its spectral representation to the spectral representations of the musical works in the plurality;

(b)   compares its rhythmic beat representation to the rhythmic beat representations of the musical works in the plurality; and

(c)   sums, including respective weighting of results of the comparison in (a) and (b), the summed results providing an indication of which musical works in the plurality are similar to the given musical work of interest.

46.   A computer data signal embodied in a carrier wave for determining similarity between a plurality of musical works, comprising:

program code for obtaining digitized audio files of the plurality, and for each musical work in the obtained digitized audio files, the program code forms (i) a spectral representation from the corresponding audio file and (ii) a rhythmic beat representation from the corresponding audio file; and

for a given musical work of interest, program code that:

(a)   compares its spectral representation to the spectral representations of the musical works in the plurality;

(b)   compares its rhythmic beat representation to the rhythmic beat representations of the musical works in the plurality; and

(c)   sums, including respective weighting of results of the comparison in (a) and (b), the summed results providing an   indication of which musical works in the plurality are similar to   the given musical work of interest.

47.   A computer system comprising:

a processor;

a memory system connected to the processor; and

a computer program, in the memory, which determines similarity between a plurality of musical works by:

obtaining respective digitized audio files of the plurality of musical works;

for each musical work, forming (i) a spectral representation from the corresponding audio file and (ii) a rhythmic beat representation from the corresponding audio file;

5   for a given musical work of interest:

(a)   comparing its spectral representation to the spectral representations of the musical works in the plurality;

(b)   comparing its rhythmic beat representation to the rhythmic beat representations of the musical works in the plurality; and

10   (c)   summing, including respective weighting of results of the comparisons in (a) and (b), said summed results providing an indication of which musical works in the plurality are similar to the given musical work of interest.

48.   A system for determining similarity between a plurality of musical works, the system

15   comprising:

means for obtaining respective digitized files of the plurality of musical works;

for each musical work, means for forming (i) a spectral representation from the corresponding audio file and (ii) a rhythmic beat representation from the corresponding audio file;

20   for a given musical work of interest:

(a)   means for comparing its spectral representation to the spectral representations of the musical works in the plurality;

(b)   means for comparing its rhythmic beat representation to the rhythmic beat representations of the musical works in the plurality; and

25   (c)   means for summing, including respective weighting of results of the comparisons in (a) and (b), the summed results providing an indication of which musical works in the plurality are similar to the given musical work of interest.